Discrete maximum principle for a space-time least squares formulation of the transport equation with finite element.

Kadidja Benmansour, Elie Bretin, Loic Piffet, Jérôme Pousin

Discrete Maximum principle for a space-time least squares formulation of the transport equation with finite element.

K. Benmansour; E. Bretin; L. Piffet and J. Pousin

Preprint (May 2013)

DISCRETE MAXIMUM PRINCIPLE FOR A SPACE-TIME INTEGRATED LEAST SQUARES FORMULATION OF THE TRANSPORT EQUATION IN FINITE ELEMENT.

K. BENMANSOUR; E. BRETIN; L. PIFFET AND J. POUSIN*

Abstract. Finite element methods are known to produce spurious oscillations when the transport equation is solved. In this paper, a variational formulation for the transport equation is proposed, and by introducing a positivity constraint combined with a penalization of the total variation of the solution, a discrete maximum principle is verified for lagrange first order finite element methods. Moreover, the oscillations are cancelled.

Key words. transport equation, variational formulation, discrete maximum principle, constrained optimization problem

1. Introduction. This paper is concerned with the discrete maximum principle in a finite element context for a space-time integrated least squares formulation of the transport equation. The transport equation (called optical flow equation), augmented with some constraints, is widely use in imaging processes see [3] for example or in computational anatomy [23]. A space-time integrated least squares formulation is well adapted in such a context since the problem is formulated as an optimization problem which allows to account for a large variety of constraints. Moreover, this formulation is known not to add diffusion in the orthogonal directions of integral curves. The least squares method has been studied with Galerkin discontinuous finite element for the transport equation in [15] and is also widely used to solve partial differential equations, see [16] and [17] for elasticity and fluid mechanics problems, and [11] for some applications in a finite element context.
The maximum principle for transport equation is sufficient to garantee positivity, monotonicity and non increasing total variation of the solution. An important feature of the finite element method for simulating transport phenomena is its inability to satisfy the maximum principle on general meshes for of the standard Galerkin formulation. This deficiency manifests itself in spurious oscillations (undershoots and overshoots). That is well known, from a long time in fluid dynamics literature.
Many remedies are available for finite differences techniques such as: the slope limiting; the flux correction, which have been recently extended to the finite element method in [19] The partial differential operator in time of the transport equation is separated in order to take advantage of the flow properties of a differential equation, and the partial differential operator in space is treated in a specific way. Remark that this technique does not extend to the formulation considered here, a space-time least-squares formulation. Some methods for handling the Dirichlet boundary condition with a least squares formulation are investigated: extension of the Dirichlet's condition; penalization of the Dirichlet's condition; Nitche's method for the Dirichlet's condition. Then some projection methods on the set of non negative functions are also investigated. But, unfortunately, with these methods some oscillations still persist. The key idea for recovering a discrete maximum principle we propose in this paper, is to transform the problem in an inequality constrained optimization problem, in order to handle the total variation and the non negativity of the solution.
In section 2 a description of the problem is given and functional spaces are introduced. In section 3 a variational formulation of the problem is given, and a weak maximum principle is proved by using a penalization technique. In section 4 a Galerkin formulation with Lagrange finite element $\mathbb{Q}^1$ is introduced, and some numerical experiments are presented for a time

*Université de Lyon ICJ UMR-CNRS 5208; 20 av. Albert Einstein; 69100 Villeurbanne, France.. e-mail: firstname.lastname@insa-lyon.fr

marching strategy. Penalized and Nitche formulations are studied numerically for handling the Dirichlet boundary condition. In section 5 Some projection techniques on the set of non negative functions are proposed and analyzed. Finally in section 6 a formulation which penalizes the total variation and the non negativity of the solution is introduced. An existence result is given, and it is proved that the total variation can be controlled according to the total variation of the data (initial condition). An algorithm FISTA is given for computing the solution, and numerical examples in 1D or 2D are presented, demonstrating that the spurious oscillations present in the previous numerical schemes do not exist anymore.

2. The problem description and the functional setting. Let $\Omega \subset \mathbb{R}^d$ (with $d = 1; 2; 3$) be a bounded domain with a Lipschitz boundary $\partial\Omega$ satisfying the cone property. If $T > 0$ is given, set $Q = \Omega \times]0, T[$. Consider an advection velocity $v : Q \to \mathbb{R}^d$ and $f : Q \to \mathbb{R}$ a given source term. In all this paper, the velocity v has at least the following regularity

$$v \in L^\infty(Q)^d \text{ and } \operatorname{div}(v) \in L^\infty(Q). \tag{2.1}$$

Let

$$\Gamma_- = \{x \in \partial\Omega : (\, v(x,t) \,|\, n(x) \,) < 0\}$$

where $n(x)$ is the outer normal to $\partial\Omega$ at point x. For the sake of the presentation, it is assumed that Γ_- do not depend on t.
The problem consists in finding a function $c : Q \to \mathbb{R}$ satisfying the following partial differential equation

$$\partial_t c + (\, v(x,t) \,|\, \nabla c(x,t) \,) = f(t,x) \quad \text{in } Q, \tag{2.2}$$

and the initial and inflow boundary conditions

$$c(x,0) = c_0(x) \qquad \text{for } x \text{ in } \Omega \tag{2.3}$$

$$c(x,t) = c_1(x,t) \qquad \text{for } x \text{ on } \Gamma_-. \tag{2.4}$$

As usual, when c_1, c_0, and v are sufficiently regular, changing the source term f if necessary, one can assume that $c_1 = 0$ on Γ_-, and $c_0 = 0$ on Ω. A similar result will be given later, using a suitable trace theorem. In what follows, the functional setting for a variational formulation of the problem (2.2-2.4) will be settled, (see also [5, 6, 7]). Moreover a trace operator is recalled in this context.
For $v \in L^\infty(Q)^d$, with $\operatorname{div}(v) \in L^\infty(Q)$, define

$$\widetilde{v} = (1, v_1, v_2, \ldots, v_d)^t \in L^\infty(Q)^{d+1}$$

and for a sufficiently regular function φ defined on Q, set

$$\widetilde{\nabla}\varphi = \left(\frac{\partial\varphi}{\partial t}, \frac{\partial\varphi}{\partial x_1}, \frac{\partial\varphi}{\partial x_2}, \ldots, \frac{\partial\varphi}{\partial x_d}\right)^t,$$

and $\widetilde{n}$ denotes the outward unit vector on ∂Q. Finaly, the Euclidean inner product will be denoted by $(\,.\,|\,.\,)$. The following theorem is proved in [13].

THEOREM 2.1. *Under the assumption $v \in L^\infty(Q)^d$, and $\operatorname{div}(v) \in L^\infty(Q)$, the normal trace of v, $(\tilde{v} \mid \tilde{n})$ is in $L^\infty(\partial Q)$.*

Let now

$$\begin{aligned} \partial Q_- &= \{(x,t) \in \partial Q, (\,\widetilde{u} \,|\, \widetilde{n}\,) < 0\} \\ &= \Gamma_- \times (0,T) \,\cup\, \Omega \times \{0\}, \end{aligned}$$

and set

$$c_b(x,t) = \begin{cases} c_0(x) & \text{if} \quad (x,t) \in \Omega \times \{0\} \\ c_1(t,x) & \text{if} \quad (x,t) \in \Gamma_- \times (0,T). \end{cases} \tag{2.5}$$

Here it is assumed that $c_b \in L^2(\partial Q_-)$. For $\varphi \in \mathcal{D}(\overline{Q})$, consider the norm

$$\|\varphi\|_{H(v,Q)} = \left(\|\varphi\|^2_{L^2(Q)} \; + \; \left\| \left(\widetilde{v} \,|\, \widetilde{\nabla} \varphi \right) \right\|^2_{L^2(Q)} + \int_{\partial Q_-} |\, (\,\widetilde{v}\,|\,\widetilde{n}\,)\,| \varphi^2 \, d\widetilde{\sigma} \right)^{1/2},$$

(see also [5, 6, 7, 9]) and then define the space $H(v,Q)$ as the closure of $\mathcal{D}(\overline{Q})$ for this norm:

$$H(v,Q) = \overline{\mathcal{D}(\overline{Q})}^{H(v,Q)}$$

If v is regular enough, it can be seen that

$$H(v,Q) = \left\{ \rho \in L^2(Q), \left(\widetilde{v} \,|\, \widetilde{\nabla} \rho \right) \in L^2(Q), \rho|_{\partial Q_-} \in L^2(\partial Q_-, |\, (\,\widetilde{v}\,|\,\widetilde{n}\,)\,|\, d\widetilde{\sigma}) \right\}$$

(see e.g. [20, 18]). Let us recall a trace result for functions belonging to $H(u,Q)$ (see [21]).
PROPOSITION 2.2. *Under the assumption $v \in L^\infty(Q)^d$, and div $(v) \in L^\infty(Q)$ there exists a linear continuous trace operator*

$$\begin{aligned} \gamma_{\widetilde{n}} : \; H(v,Q) &\longrightarrow L^2(\partial Q, |\, (\,\widetilde{v}\,|\,\widetilde{n}\,)\,| d\widetilde{\sigma}) \\ \varphi \; &\mapsto \; \gamma_{\widetilde{n}} \varphi = \varphi_{|\partial Q}, \end{aligned}$$

which can be localized as:

$$\begin{aligned} \gamma_{\widetilde{n}_\pm} : \; H(v,Q) &\longrightarrow L^2(\partial Q_\pm, |\, (\,\widetilde{v}\,|\,\widetilde{n}\,)\,| d\widetilde{\sigma}) \\ \varphi \; &\mapsto \; \gamma_{\widetilde{n}_\pm} \varphi = \varphi_{|\partial Q_\pm}. \end{aligned}$$

Finally define the space

$$\begin{aligned} H_0 = H_0(u, Q, \partial Q_-) &= \{ \rho \in H(u,Q), \; \rho = 0 \text{ on } \partial Q_- \} \\ &= H(u,Q) \cap \text{ Ker } \gamma_{\widetilde{n}_-}. \end{aligned}$$

We now recall an extension of the *curved Poincaré inequality* obtained in [5, 6].

THEOREM 2.3. *If $v \in L^\infty(Q)^d$ and div $(v) \in L^\infty(Q)$, the semi-norm on $H(u,Q)$ defined by*

$$|\rho|_{1,v} = \left(\int_Q \left(\widetilde{v} \,|\, \widetilde{\nabla} \rho \right)^2 dx \; dt + \int_{\partial Q_-} |\, (\,\widetilde{v}\,|\,\widetilde{n}\,)\,| \rho^2 \, d\widetilde{\sigma} \right)^{1/2} \tag{2.6}$$

is a norm, equivalent to the norm given on $H(v,Q)$.
Henceforth the space $H(v,Q)$ is equipped with the norm $|\varphi|_{1,v}$.

REMARK 1. *a) Using the above result, if $c_b = 0$, the semi-norm*

$$|\rho|_{1,v} = \left(\int_Q \left(\widetilde{v} \,|\, \widetilde{\nabla} \rho \right)^2 dx \; dt \right)^{1/2}$$

in a norm on H_0 which is equivalent to the usual norm on $H(v,Q)$.
Let us end this section with a least squares formulation in $L^2(Q))$. A space-time least squares solution of equation (2.2) corresponds to a minimizer in
$\{\varphi \in H(v,Q); \gamma_{\widetilde{n}_-}(\varphi) - c_b = 0\}$ of the following convex, $H(v,Q)$-coercive functional

$$J(c) = \frac{1}{2}\left(\int_Q \left(\left(\widetilde{v} \,|\, \widetilde{\nabla} c\right) - f\right)^2 dx\, dt - \int_{\partial Q_-} c^2 \left(\widetilde{u} \,|\, \widetilde{n}\right) d\widetilde{\sigma}\right).$$

The Gâteau derivative of J is

$$DJ(c)\varphi = \int_Q \left(\left(\widetilde{v} \,|\, \widetilde{\nabla} c\right) - f\right)\ \left(\widetilde{v} \,|\, \widetilde{\nabla} \varphi\right) dx\, dt - \int_{\partial Q_-} c\varphi \left(\widetilde{v} \,|\, \widetilde{n}\right) d\widetilde{\sigma}.$$

So a sufficient condition to get the least squares solution of (2.2 - 2.4) is the following *weak formulation*: If $c_b \in L^2(\partial Q_-)$, find $c \in H(v,Q)$ such that

$$\int_Q \left(\widetilde{v} \,|\, \widetilde{\nabla} c\right) \cdot \left(\widetilde{v} \,|\, \widetilde{\nabla} \varphi\right) dx\, dt = \int_Q f \cdot \left(\widetilde{v} \,|\, \widetilde{\nabla} \varphi\right) dx\, dt; \tag{2.7}$$
$$\gamma_{\widetilde{n}_-}(c) = c_b$$

for all $\varphi \in H_0$ (see [5, 6, 7, 9, 14, **?**]).

3. A weak maximum principle for the space-time least squares formulation. This section is devoted to the study of equation (2.7). More precisely, an existence and uniqueness theorem for the solution of equation (2.7) is given. Then a weak maximum principle is deduced. First, let us give a penalized formulation of the space-time least squares, useful for some L^∞ estimate [21].
LEMMA 3.1. *If $c_b \in G_-$, let c^m be the solution of*

$$\int_Q \left(\widetilde{v} \,|\, \widetilde{\nabla} c^m\right)\left(\widetilde{v} \,|\, \widetilde{\nabla} \varphi\right) dx\, dt - m\int_{\partial Q_-} (c^m - c_b)\cdot \varphi \left(\widetilde{v} \,|\, \widetilde{n}\right) d\widetilde{\sigma} = \int_Q f \cdot \left(\widetilde{v} \,|\, \widetilde{\nabla} \varphi\right) dx\, dt, \tag{3.1}$$

$\forall \varphi \in H(v,Q)$. *There is a subsequence of c^m which weakly converges in $H(v,Q)$ to the solution c of (2.7).*
Later we will use the following versions of Stampacchia's theorem (see [**?**, **?**]).

THEOREM 3.2. *Let $\rho \in H(u,Q)$, then*

$$\left(\widetilde{v} \,|\, \widetilde{\nabla} \rho\right) = 0 \ \textit{ a.e. on the set } \ \{(x,t) \in Q; \rho(x.t) = 0\} \tag{3.2}$$

THEOREM 3.3. *Let $g : \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function.*
a) If $\rho \in H(u,Q)$, then $g(\rho) \in H(u,Q)$.
b) If g is differentiable except at a finite number of points, say $\{z_1,\ldots,z_n\}$, then

$$\left(\widetilde{v} \,|\, \widetilde{\nabla} g(\rho)\right) = \begin{cases} g'(\rho)\ \left(\widetilde{v} \,|\, \widetilde{\nabla} \rho\right) & \textit{if } \rho(x,t) \notin \{z_1,\ldots,z_n\} \\ 0 \ \textit{ elsewere.} & \end{cases} \tag{3.3}$$

Now reduce the problem (2.2)-(2.4) to an homogeneous Dirichlet problem on ∂Q_-. For $c_b \in L^2(\partial Q_-)$, let $C_b \in H(u,Q)$ be such that $\gamma_{\widetilde{n}_-}(C_b) = c_b$. Then $\rho = c - C_b$ is the unique solution of

$$\int_Q \left(\widetilde{v} \,|\, \widetilde{\nabla}\rho \right) \cdot \left(\widetilde{v} \,|\, \widetilde{\nabla}\psi \right) \, dx \, dt = \int_Q \left(f - \left(\widetilde{v} \,|\, \widetilde{\nabla} C_b \right) \right) \cdot \left(\widetilde{v} \,|\, \widetilde{\nabla}\psi \right) \, dx \, dt \tag{3.4}$$

for all $\psi \in H_0$. Moreover the solution of problem (3.4) is equivalent to the solution of (2.2).

THEOREM 3.4. *For a fixed $v \in L^\infty(Q)^d$ with div $(v) \in L^\infty(Q)$, and $c_b \in L^2(\partial Q_-)$, and $f \in L^2(Q)$, the problem (3.4) has a unique solution. Moreover*

$$|\rho|_{1,v} = \left\| \left(\widetilde{v} \,|\, \widetilde{\nabla}\rho \right) \right\|_{L^2(Q)} \leq \|f\|_{L^2(Q)} + \left\| \left(\widetilde{v} \,|\, \widetilde{\nabla} C_b \right) \right\|_{L^2(Q)},$$

and the function $c = \rho + C_b$ is the space-time least squares solution *of (2.2).*

Proof. This assertion is a consequence of the Curved Poincaré inequality (theorem 2.3) and of the Lax-Milgram theorem (see also [5, 6]). ∎

REMARK 2. *For the numerical solution of equation (2.7), a time marching approach can be used to avoid the consideration of the all of Q (see e.g. [9, 14,* **?***]).*

COROLLARY 3.5. *The solution c of equation (2.7) belongs to the space*

$$X = L^2(Q) \cap L^2(\partial Q_+, (\,\widetilde{u} \,|\, \widetilde{n}\,)\, d\widetilde{\sigma})$$

equipped with the norm $|||c|||$.

The following theorem is a weak maximum principle for the solution of problem (2.7).

THEOREM 3.6. *Assume that the function $f = 0$ in equation (2.7) and that the function $c_b \in L^\infty(\partial Q_-)$. Then the solution of equation (2.7) satisfies*

$$\inf c_b \leq c \leq \sup c_b.$$

Proof. Let c^m be the sequence of solutions to the penalized formulation given in Lemma 3.1. Then

$$\int_Q \left(\widetilde{v} \,|\, \widetilde{\nabla} c^m \right) \cdot \left(\widetilde{v} \,|\, \widetilde{\nabla}\varphi \right) \, dx \, dt - m \int_{\partial Q_-} c^m \cdot \varphi \,(\,\widetilde{v} \,|\, \widetilde{n}\,)\, d\widetilde{\sigma} = - m \int_{\partial Q_-} c_b \cdot \varphi \,(\,\widetilde{u} \,|\, \widetilde{n}\,)\, d\widetilde{\sigma}$$

for all $\varphi \in H(v,Q)$. Set

$$M = \sup_{\partial Q_-} c_b$$

and according to theorem 3.3 put

$$\varphi = (c^m - M)^+, \; Q_1 = \left\{ (x,t) \in \overline{Q}, c^m - M > 0 \right\}, \; \Sigma_1 = \partial Q_- \cap Q_1$$

Then, from theorem 3.2,

$$\int_{Q_1} \left(\widetilde{v} \,|\, \widetilde{\nabla} c^m \right) \cdot \left(\widetilde{v} \,|\, \widetilde{\nabla} (c^m - M) \right) \, dx \, dt - m \int_{\Sigma_1} c^m \cdot (c^m - M) \, (\, \widetilde{v} \,|\, \widetilde{n} \,) \, d\widetilde{\sigma} =$$
$$- \, m \int_{\Sigma_1} c_b \cdot (c^m - M) \, (\, \widetilde{v} \,|\, \widetilde{n} \,) \, d\widetilde{\sigma}.$$

We have,

$$\int_{Q_1} \left(\widetilde{v} \,|\, \widetilde{\nabla} (c^m - M) \right)^2 \, dx \, dt - m \int_{\Sigma_1} ((c^m - M))^2 \, (\, \widetilde{v} \,|\, \widetilde{n} \,) \, d\widetilde{\sigma} =$$
$$- \, m \int_{\Sigma_1} (c_b - M) \cdot (c^m - M) \, (\, \widetilde{v} \,|\, \widetilde{n} \,) \, d\widetilde{\sigma} \ \leq 0.$$

Hence, the set Q_1 has a zero measure, so $c^m \leq M$. We show in the same way that $c^m \geq \inf \ c_b$. Finally the conclusion holds for the weak limit c of a subsequence of the sequence c^m. ■

□

In the following, it is assumed that at least, the following regularity holds true:
H) $v \in C^0(\overline{Q})$, $c_b \in C^0(\partial Q_-)$.

4. Finite element approximation . Let $\{\varphi_1, \varphi_2, \dots, \varphi_N\}$ be a basis of finite element subspace $V_h \subset H_0$, obtained, for example,with a rectangular mesh $\mathcal{T}_h$ of the domain Q, with first order $\mathbb{Q}_1$ Lagrange quadrangular finite element:

$$Q = \bigcup_{\mathcal{T}_h} T, \ \ V_h = \left\{ \varphi \in C^0(\overline{Q}) \mid \varphi_{|T} \in \mathbb{Q}^1(T) \right\}. \tag{4.1}$$

Define the bilinear symmetric form $a(\cdot, \cdot) \, V_h \times V_h \to \mathbb{R}$ by

$$a(\psi_h, \varphi_h) = \int_Q \left(\widetilde{v} \,|\, \widetilde{\nabla} \psi_h \right) \left(\widetilde{v} \,|\, \widetilde{\nabla} \varphi_h \right) \, dxdt.$$

An approximation of problem (3.4), where the right hand side $f - \left(\widetilde{v} \,|\, \widetilde{\nabla} C_b \right)$ is still denoted by f for keeping the notations as simple as possible, consists in finding $c_h \in V_h$ such that

$$a(\varphi_h, c_h) = \int_Q f \left(\widetilde{v} \,|\, \widetilde{\nabla} \varphi_h \right) \, dx \, dt; \tag{4.2}$$

for all $\varphi_h \in V_h$, where

$$c_h = \sum_{j=1}^{N} \varphi_j(t, x) \cdot c_j.$$

With these notations, equation (4.2) becomes

$$\sum_{j=1}^{N} c_j \int_Q \left(\widetilde{v} \,|\, \widetilde{\nabla} \varphi_j \right) \left(\widetilde{v} \,|\, \widetilde{\nabla} \varphi_i \right) \, dx \, dt = \int_Q f \left(\widetilde{v} \,|\, \widetilde{\nabla} \varphi_i \right) \, dx \, dt; \tag{4.3}$$

for all $i = 1, \ldots, N$. Set

$$a_{ij} = \int_Q \left(\widetilde{v} \,|\, \widetilde{\nabla}\varphi_j \right) \left(\widetilde{v} \,|\, \widetilde{\nabla}\varphi_i \right) \, dx \, dt, 1 \leq i, j \leq N;$$

and

$$b_i = \int_Q f \left(\widetilde{v} \,|\, \widetilde{\nabla}\varphi_i \right) \, dx \, dt; 1 \leq i \leq N;$$

The coefficients, a_{ij}, b_i, are computed in the standard way.
If $A = (a_{ij})_{1 \leq i,j \leq N}$, $B = (b_i)_{1 \leq i \leq N}$ and $C = (c_i)_{1 \leq i \leq N}$, then the solution of the linear system

$$AC = B \tag{4.4}$$

is the solution of problem (4.2).
The method we consider is a marching technique, that is to say the solution is computed time slice by time slice. Such a strategy is possible since $\widetilde{v}_1 = 1$, then the integral curves associated to $\widetilde{v}$ are increasing with respect to time. At each time step we solve a "local time" problem where the initial condition is c_h at the current time step and the unknown is c_h at the next time step. The stiffness matrix of the system is calculated for a single slice of finite elements of width $\triangle t$. Here, the solution c_h comprising only the solutions computed at time t (taken as an initial condition or boundary condition) and at time $t + \triangle t$. The system is solved and we proceed step by step until the final time is reached. Assume the domain $Q = \Omega \times \triangle t$, as mentioned above, and let $V_h \subset H(v, Q)$ be a first order $\mathbb{Q}^1$ Lagrange finite element subspace. The vector C is decomposed as follows: C_- containing only values on nodes at time t and the unknowns $\widetilde{C}$ at time $t+\triangle t$, such as degrees of freedom belonging to ∂Q_- have the lower labels. We have:

$$AC = \begin{pmatrix} M & N \\ P & Q \end{pmatrix} \begin{pmatrix} C_- \\ \widetilde{C} \end{pmatrix} = \begin{pmatrix} B_- \\ \widetilde{B} \end{pmatrix} \tag{4.5}$$

The solution $\widetilde{C}$ at time step $t + \triangle t$ is obtained by solving the reduced system:

$$Q\widetilde{C} = \widetilde{B} - PC_-. \tag{4.6}$$

4.1. Numerical experiments for the 1D time marching scheme. In this section, we give a numerical example with the transport equation with $v = 1$ and with only a non zero initial condition, taking the initial condition as follows $c_0(x,t) = \frac{1}{2}(1 - tanh(100x - 50))$. The number of elements is $n = 80$ and the time step is $\triangle t = 1/80$. In the following figures 4.1, the solution is represented after 1 and 20 time steps.
In figure 4.1 the solution exhibits oscillations as time elapses. Despite these oscillations, the method converges when n goes to infinity, as it is shown in figure 4.2.
In the next figure 4.3, a penalized formulation as in section 3 with a penalization parameter equal to 80 has been implemented, and the solution is presented after one and 20 time steps, keeping the same values for the other parameters.

4.1.1. A Nitche's method. First, let us mention that least squares finite element methods for non conforming (discontinuous) finite element have been investigated in [15]. It is

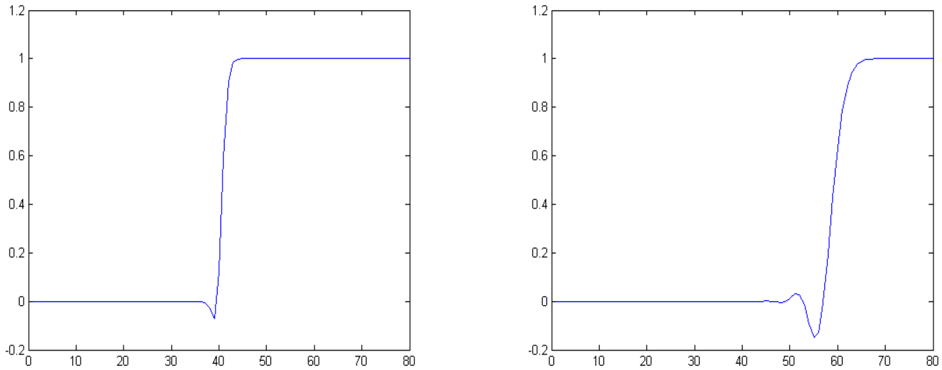


FIG. 4.1. *(a) Solution for one time step; (b) Solution for 20 time steps.*

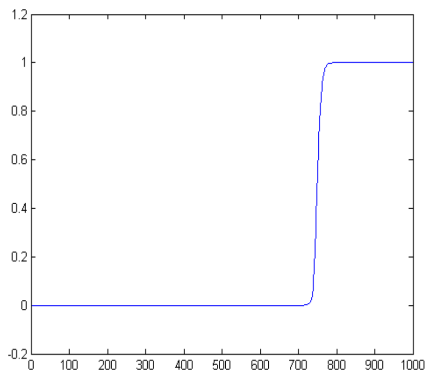


FIG. 4.2. *convergence of the solution for n=1000.*

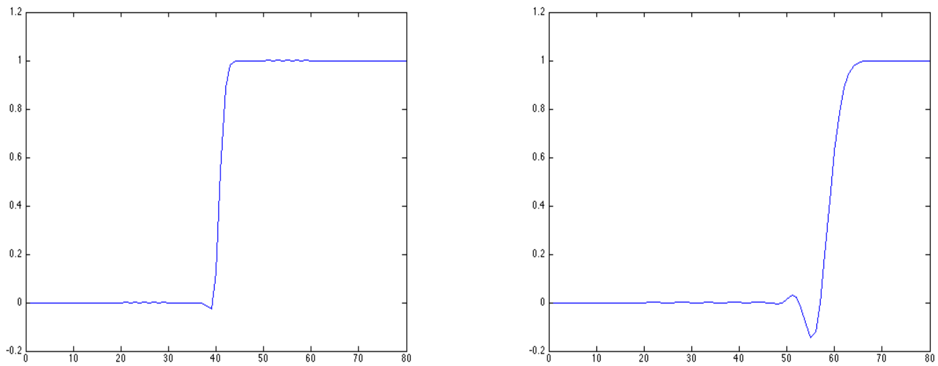
reported that spurious oscillations exist for that method when polynomial degree of approximation ranging from 1 to 4 (see [15] P. 50). In what follows, a Nitche's formulation of the problem is proposed, since it is very close to the penalized formulation presented in section 3. Starting from the equations (2.7), denote $(\widetilde{v}\otimes\widetilde{v})$ the matrix the coefficients of which are $\widetilde{v}_i\widetilde{v}_j$ for $1\leq i,j\leq 2$ and accounting for

$$(\,a\,|\,b\,)\,(\,a\,|\,c\,)=(\,(a\otimes a)b\,|\,c\,)$$

the following partial differential equations are obtained:

$$\left\{\begin{array}{lr} -\widetilde{\mathrm{div}}((\widetilde{v}\otimes\widetilde{v})\widetilde{\nabla}c)\,dx\;dt=-\widetilde{\mathrm{div}}(\widetilde{v}f) & \text{in } Q\\ c=c_b; & \text{on } \partial Q_-;\\ \Big(\,(\widetilde{v}\otimes\widetilde{v})\widetilde{\nabla}c\,|\,\widetilde{n}\,\Big)=f\,(\,\widetilde{v}\,|\,\widetilde{n}\,) & \text{on } \partial Q_+. \end{array}\right. \tag{4.7}$$

Since we are interested in a discrete maximum principle set $f=0$. A Nitche's method for the Dirichlet's condition on the boundary ∂Q_- reads: find $c_h\in V_h\subset H(v,Q)$ the dimension of



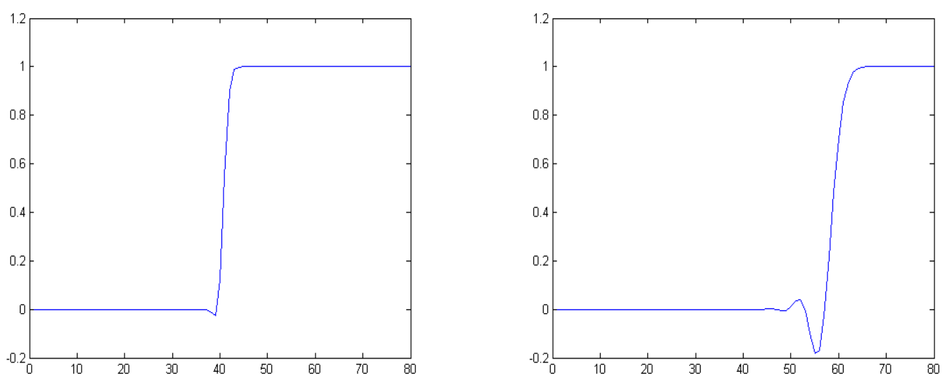
(a) (b)

FIG. 4.3. *(a) Solution of a penalized formulation for one time step; (b) Solution of a penalized formulation for 20 time steps.*

which is N solution to

$$
\begin{aligned}
&\textstyle\sum_{j=1}^{N} c_j \int_Q \left(\widetilde{v} \,|\, \widetilde{\nabla}\varphi_j \right) \left(\widetilde{v} \,|\, \widetilde{\nabla}\varphi_i \right) \, dx \; dt + 1/\gamma \sum_{j=1}^{N} (c_j - u_{c_{b_j}}) \int_{\partial Q_-} \varphi_j \varphi_i \, d\sigma \\
&- \textstyle\sum_{j=1}^{N} c_j \int_{\partial Q_-} \left(\widetilde{v} \,|\, \widetilde{\nabla}\varphi_j \right) \varphi_i \,(\widetilde{v} \,|\, \widetilde{n}) \; d\sigma \qquad\qquad = 0
\end{aligned}
$$

for all $\varphi_i \in V_h$ and whatever γ is. In the next figure 4.1.1, the solution to the Nitche's formulation is presented with $n = 80$ and $\triangle t = 1/100$ and $\gamma = \frac{n}{20}$.



(a) (b)

FIG. 4.4. *(a) Solution of Nitche's formulation for one time step; (b) Solution of Nitche's formulation formulation for 20 time steps.*

5. Projection methods and Maximum principle. Remark that the basis functions φ_k, $1 \geq k \geq N$ are non-negative. Thus, we define the convex subset K_h as following :

$$K_h = \left\{ \sum_{k=1}^{N} \alpha_k \varphi_k \;\middle|\; \alpha_k \in \mathbb{R}^+ \right\}. \tag{5.1}$$

In order to recover the discrete maximum principle we introduce a projection step onto the cone of non negative functions K_h. In what follows, an orthogonal projection with the inner product induced by the bilinear form $a(\cdot,\cdot)$ is considered, then a nodal projection is introduced.

5.1. A H_0 projection onto the cone of positive functions. At first assume f to be regular and A to be a M-matrix. Define c_{h_p} the projection of c_h onto K_h with respect to the inner product $a(\cdot,\cdot)$. We have for all $\psi \in K_h$:

$$a(c_h - c_{h_p}, \psi - c_{h_p}) \leq 0.$$

Set $\varphi = \psi + c_{h_p}$, we have

$$a(c_h - c_{h_p}, \varphi) \leq 0 \,\forall \varphi \in K_h$$

Now, let us specify c_{h_p}. Set $g = \left(\widetilde{v} \,|\, \widetilde{\nabla} f\right)^- ; h = f^+$ (where $z^\pm$ denotes the positive or negative part of z) and define c_{h_p} as to be solution to

$$a(c_{h_p}, \varphi) = \int_Q g\varphi \, dx\, dt + \int_{\partial Q_+} h\varphi \left(\widetilde{v} \,|\, \widetilde{n}\right) \, ds. \tag{5.2}$$

Integrate by parts the right end side of Eqaution (4.2),

$$\int_Q f \left(\widetilde{v} \,|\, \widetilde{\nabla} \varphi_h\right) \, dx\, dt = - \int_Q \left(\widetilde{v} \,|\, \widetilde{\nabla} f\right) \varphi \, dx\, dt + \int_{\partial Q_+} f\varphi \left(\widetilde{v} \,|\, \widetilde{n}\right) \, ds;$$

we have

$$a(c_h - c_{h_p}, \varphi) = \int_Q (-\left(\widetilde{v} \,|\, \widetilde{\nabla} f\right) - g)\varphi \, dx\, dt + \int_{\partial Q_+} (f - h)\varphi \left(\widetilde{v} \,|\, \widetilde{n}\right) \, ds.$$

Thus

$$a(c_h - c_{h_p}, \varphi) = \int_Q - \left(\widetilde{v} \,|\, \widetilde{\nabla} f\right)^+ \varphi \, dx\, dt - \int_{\partial Q_+} f^- \varphi \left(\widetilde{v} \,|\, \widetilde{n}\right) \, ds \leq 0;$$

for all $\varphi \in K_h$. For c_{h_p} to be in K_h it is required A to be a M-matrix since the right end side is non negative.
In the example we considered, the initial condition $c_0(x,t) = \frac{1}{2}(1 - tanh(100x - 50))$, and we had a constant velocity equats to 1, so in this case the matrix A is a M-matrix. By using the H_0 projection onto the cone K_h the oscillations (”under shooting”) are reduced, but do not disappear as it is shown in Figure 5.1. The persisting discrepancy is due to the poor approximation of the positive or negative part with finite element.

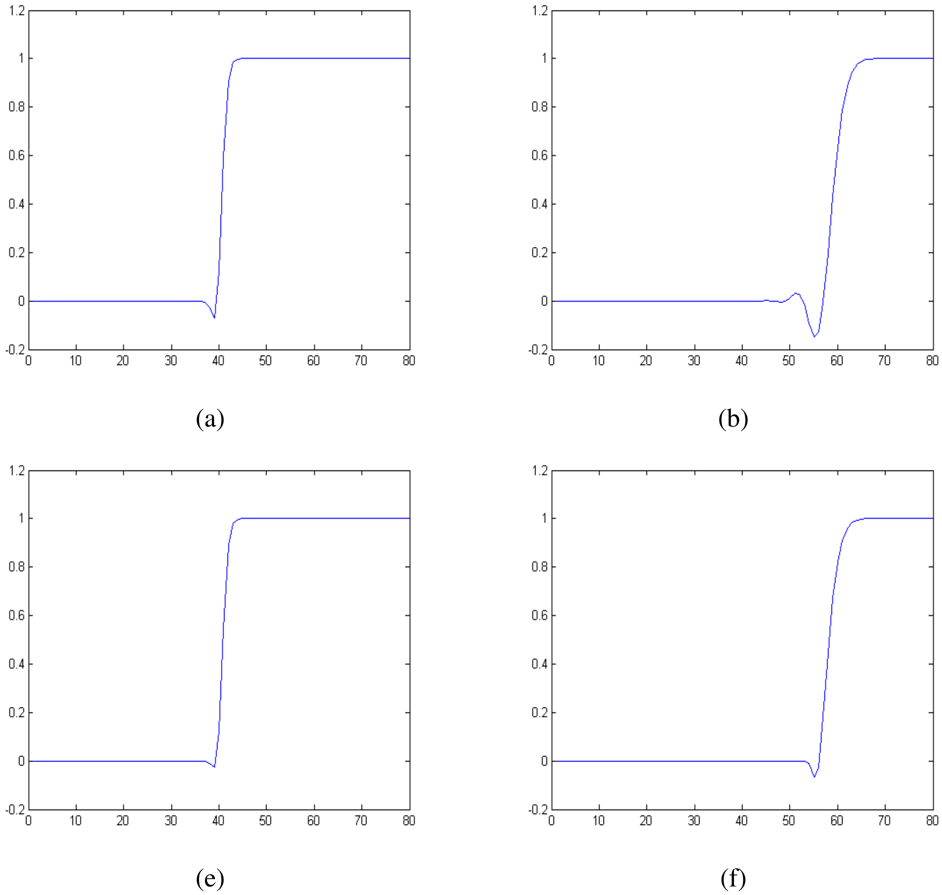


FIG. 5.1. *(a) Time marching solution for one time step; (b) for 20 time steps.(e) Time marching solution to H_0 projection problem (5.2) for one time step; (f) for 20 time steps*

5.2. Generalized Lagrange multiplier method. Let us rephrase the problem (4.4) as a constrained optimization problem in $\mathbb{R}^N$. Find C_p satisfying :

$$\begin{cases} AC_p = B \\ C_p \geq 0. \end{cases} \tag{5.3}$$

Express the previous problem as the following minimization problem on $\mathbb{R}^N_+$:

$$C_p = \underset{X \in \mathbb{R}^N_+}{\text{Agmin}} \quad \frac{1}{2} X^t A X - X^t B. \tag{5.4}$$

The problem is well posed, and by using the complementarity conditions [22]

$$0 \leq (AC_p - F); C_p \perp (AC_p - F) \tag{5.5}$$

problem (5.4) can be be computed by using the following generalized Lagrange multiplier method for $0 < r$ fixed.

$$\begin{cases} AC_p = B + \Lambda \\ \Lambda = (\Lambda - rC_p)^+ \end{cases} \tag{5.6}$$

where, the positive part of a vector denotes the positive part of its components. An iterative algorithm is proposed for solving the problem (5.6). Set $C_p^0 = 0_{\mathbb{R}^N}; \Lambda^0 = 0_{\mathbb{R}^N}$, then compute:

$$\left\{ \begin{array}{l} AC_p^{k+1} = B + \Lambda^{k+1} \\ \Lambda^{k+1} = (\Lambda^k - rC_p^k)^+ \end{array} \right. \tag{5.7}$$

Now the convergence of the iterative procedure is proved. We have:
LEMMA 5.1. *Let μ_1 the first eigenvalue of the positive definite matrix A. For all r verifying $0 < r < 2\mu_1$ the algorithm 5.7 converges.*
Proof.

$$\left\{ \begin{array}{c} A(C_p^{k+1} - C_p^k) = \Lambda^{k+1} - \Lambda^k \\ \Lambda^{k+1} - \Lambda^k = (\Lambda^k - rC_p^k)^+ - (\Lambda^{k-1} - rC_p^{k-1})^+. \end{array} \right. \tag{5.8}$$

Since z^+ is a 1-lipschitzian function we deduce:

$$\left\|\Lambda^{k+1} - \Lambda^k\right\|^2 \leq \left\|\Lambda^k - \Lambda^{k-1}\right\|^2 - 2r\left(\Lambda^k - \Lambda^{k-1} \,|\, C_p^k - C_p^{k-1}\right) + r^2\left\|C_p^k - C_p^{k-1}\right\|^2.$$

Since the matrix A is positive definite, by using the first equation of (5.8), the previous inequality becomes:

$$\begin{aligned} \left\|\Lambda^{k+1} - \Lambda^k\right\|^2 &\leq \left\|\Lambda^k - \Lambda^{k-1}\right\|^2 - 2r\left(A(C_p^k - C_p^{k-1}) \,|\, C_p^k - C_p^{k-1}\right) + r^2\left\|C_p^k - C_p^{k-1}\right\|^2 \\ &\leq \left\|\Lambda^k - \Lambda^{k-1}\right\|^2 + r(r - 2\mu_1(A))\left\|C_p^k - C_p^{k-1}\right\|^2. \end{aligned}$$

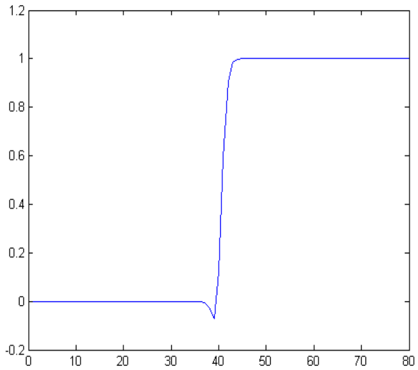
If $\left\|C_p^k - C_p^{k-1}\right\|^2 = 0$, the sequence $\{\Lambda^p\}_{p>k}$ becomes stationary and thus converges, if $\left\|C_p^k - C_p^{k-1}\right\|^2 \neq 0$ we have the existence of $\xi < 1$, such that:

$$\left\|\Lambda^{k+1} - \Lambda^k\right\|^2 < \left\|\Lambda^k - \Lambda^{k-1}\right\|^2 \quad \text{that is also expressed as}$$
$$\left\|\Lambda^{k+1} - \Lambda^k\right\|^2 \leq \xi\left\|\Lambda^k - \Lambda^{k-1}\right\|^2.$$

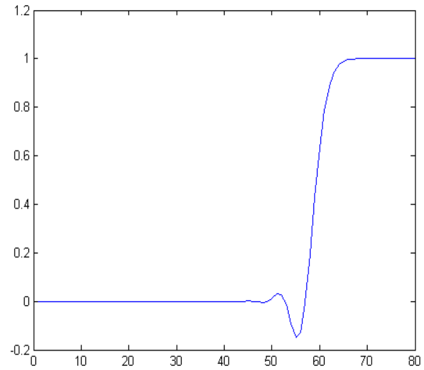
For all $q < p$ we deduce:

$$\|\Lambda^p - \Lambda^q\|^2 \leq \sum_{l=q+1}^{l=p}\left\|\Lambda^l - \Lambda^{l-1}\right\|^2 \leq \sum_{l=q+1}^{l=p}\xi^{l-1}\left\|\Lambda^1\right\|^2 \leq \xi^q\sum_{m=0}^{\infty}\xi^m\left\|\Lambda^1\right\|^2$$

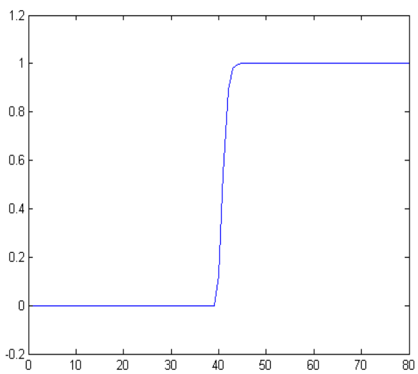
which proves that $\{\Lambda^k\}_{k\in\mathbb{N}}$ is a Cauchy's sequence.
The sequence $\{C_p^k\}_{k\in\mathbb{N}}$ is also a Cauchy's sequence, and we can take the limit in the equations.
□
Considering as before the 1D example, the obtained numerical results are presented in figure 5.2.
REMARK 3.
The generalized Lagrange multiplier method is not a projection technique, thus the maximum of the constrained problem could increase when the generalized Lagrange multiplier method is used.
If the solution is not a monotonic function, the generalized Lagrange multiplier method is not able to handle spurious oscillations as it is examplefied with the following initial condition. Set $c_0(x,t) = 1/2 + (1/2\tanh(100x - 20)) + 1/2 + (1/2\tanh(100x - 35))$. In figure 5.2, the computed solution with the generalized Lagrange multiplier method is presented after 10 time steps with the same values of parameters as before.



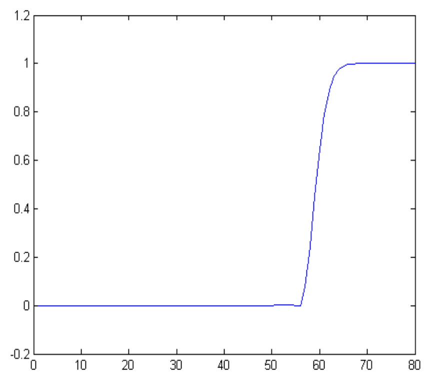
(a)



(b)



(g)



(h)

FIG. 5.2. *(a) Time marching solution for one time step; (b) for 20 time steps.(g) Time marching solution to generalized Lagrange multiplier (5.6) for one time step (h) for 20 time steps*

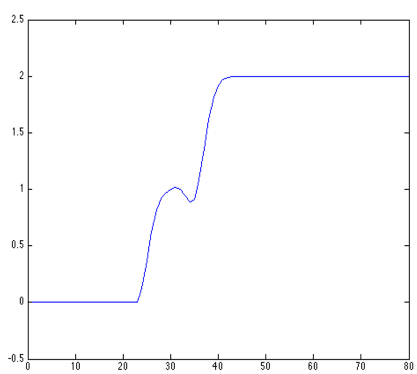


FIG. 5.3. *Time marching solution to generalized Lagrange multiplier (5.6) for 10 time steps*

6. A least squares formulation with a penalized total variation and non negativity. First, let us recall the main properties of the space of bounded variation functions $BV(Q)$ (see [1, 2, 4] for example). Introduce

$$TV(u) = \sup\left\{ \int_Q u(x)\,\mathrm{div}\,\xi(x)\,dx \mid \xi \in \mathcal{C}_c^1(Q),\ \|\xi\|_\infty \leq 1 \right\} \tag{6.1}$$

and define the space

$$BV(Q) = \{u \in L^1(Q) \mid TV(u) < +\infty\}.$$

The space $BV(Q)$, endowed with the norm $\|u\|_{BV(Q)} = \|u\|_{L^1} + TV(u)$, is a Banach space. The derivative in distributional sense of a function $u \in BV(Q)$ is a bounded Radon measure, denoted Du, and $TV(u) = \int_Q |Du|$ is the total variation of u. We have [1, 2]:

LEMMA 6.1. *The mapping $u \mapsto TV(u)$ is lower semi-continuous (denoted in short lsc) from $BV(Q)$ into $\mathbb{R}^+$ for the L^1 norm. Moreover, we have the continuous embedding $BV(Q) \subset L^q(Q)$ for every $1 \leq q \leq \frac{n}{n-1}$, which is compact if $q < \frac{n}{n-1}$.*
Note that, for $u \in W^{1,1}(Q)$, $TV(u) = \|\nabla u\|_{L^1(Q)}$.
Let K be the cone of non negative functions

$$K = \{\varphi \in H_0 \cap BV(Q),\ \varphi \geq 0 \text{ a.e.}\}, \tag{6.2}$$

and denotes I_K its indicator function.
For $\lambda \in \mathbb{R}_+$ to be fixed, consider the following optimization problem:

$$\rho_\lambda = \underset{c \in H_0 \cap BV(Q)}{\operatorname{argmin}}\ J(c) + \lambda TV(c) + I_K(c) = \underset{c \in H_0 \cap BV(Q)}{\operatorname{argmin}}\ F(c), \tag{6.3}$$

where the function J has been introduced in the first section
$J(c) = \frac{1}{2}\int_Q \left(\left(\widetilde{v}(x,t) \mid \widetilde{\nabla} c(x,t) \right) - f(t,x) \right)^2$
THEOREM 6.2. *Whatever λ non negative real is, the problem 6.3 has a unique solution.*
Proof.
Let $(c_n)_n \in H_0 \cap BV(Q)$ be a minimizing sequence, i.e.

$$\lim_{n \to +\infty} F(c_n) = \inf\{F(c) \mid c \in H_0 \cap BV(Q)\} < +\infty.$$

The sequence $(|c_n|_{1,v})_{n \in \mathbb{N}}$ is bounded, then the sequence $(c_n)_{n \in \mathbb{N}}$ weakly converges to $c^* \in H_0$, up to a subsequence. The function J is convex and l.s.c., we have

$$J(c^*) \leq \liminf_{n \to +\infty} J(c_n). \tag{6.4}$$

The semi norm $|\cdot|_{1,v}$ and the norm $\|\cdot\|_{H(v,Q)}$ are equivalent in H_0, the sequence $(c_n)_{n \in \mathbb{N}}$ is bounded in $L^2(Q)$, and $(c_n)_{n \in \mathbb{N}}$ weakly converges towards c^* in $L^2(Q)$. Furthermore, Q is bounded, the embedding of $L^2(Q)$ into $L^1(Q)$ is continuous, $(c_n)_{n \in \mathbb{N}}$ is bounded in $L^1(Q)$ as well, and therefore it is bounded in $BV(Q)$. The compact embedding of $BV(Q)$ into $L^1(Q)$ yields the strong converge in $L^1(Q)$ of a subsequence of $(c_n)_{n \in \mathbb{N}}$ towards c^*, with $c^* \in BV(Q)$. From the lemma 6.1 we have:

$$TV(c^*) \leq \liminf_{n \to +\infty} TV(c_n). \tag{6.5}$$

The subspace K is convex and closed for the L^1-norm, I_K is thus convex and l.s.c. We conclude

$$I_K(c^*) \leq \liminf_{n\to+\infty} I_K(c_n). \tag{6.6}$$

Finally we have up to a subsequence:

$$F(c^*) \leq \liminf_{n\to+\infty} F(c_n), \tag{6.7}$$

that is to say

$$F(c^*) \leq \inf \{F(c) \mid c \in H_0 \cap BV(Q)\}. \tag{6.8}$$

Uniqueness of minimal argument comes from the strict convexity of function J, (due to the L^2-norm).

□

Now let us give a technical result concerning the behavior of ρ_λ with respect to the parameter λ.

LEMMA 6.3. *Let ρ_λ be the solution to the problem 6.3. Then the application $Y : \lambda \mapsto \rho_\lambda$ from $\mathbb{R}_+$ with values in $H_0 \cap BV(Q)$ is continuous . Moreover, the application $T : \lambda \mapsto TV(\rho_\lambda)$ from $\mathbb{R}_+$ with values in $\mathbb{R}_+$ is continuous and decreasing towards zero.*

Proof.

We first show that T is a decreasing function toward zero.

Let $\lambda_2 > \lambda_1 > 0$ be given. We set $\rho_{\lambda_1} = Y(\lambda_1)$ and $\rho_{\lambda_2} = Y(\lambda_2)$. By definition of Y, we can write the two following inequalities :

$$J(\rho_{\lambda_1}) + \lambda_1 TV(\rho_{\lambda_1}) + I_K(\rho_{\lambda_1}) \leq J(\rho_{\lambda_2}) + \lambda_1 TV(\rho_{\lambda_2}) + I_K(\rho_{\lambda_2}), \tag{6.9}$$

and

$$J(\rho_{\lambda_2}) + \lambda_2 TV(\rho_{\lambda_2}) + I_K(\rho_{\lambda_2}) \leq J(\rho_{\lambda_1}) + \lambda_2 TV(\rho_{\lambda_1}) + I_K(\rho_{\lambda_1}). \tag{6.10}$$

The sum of these two inequalities gives

$$\lambda_1 TV(\rho_{\lambda_1}) + \lambda_2 TV(\rho_{\lambda_2}) \leq \lambda_1 TV(\rho_{\lambda_2}) + \lambda_2 TV(\rho_{\lambda_1}), \tag{6.11}$$

then

$$(\lambda_2 - \lambda_1)(TV(\rho_{\lambda_2}) - TV(\rho_{\lambda_1})) \leq 0, \tag{6.12}$$

that is to say

$$TV(\rho_{\lambda_2}) - TV(\rho_{\lambda_1}) \leq 0. \tag{6.13}$$

The decreasing of T follows.

Let $\lambda > 0$ be given. If we set $\rho_\lambda = Y(\lambda)$, we have, for all $\rho \in H_0 \cap BV(Q)$

$$J(\rho_\lambda) + \lambda TV(\rho_\lambda) + I_K(\rho_\lambda) \leq J(\rho) + \lambda TV(\rho) + I_K(\rho). \tag{6.14}$$

In particular,

$$J(\rho_\lambda) + \lambda TV(\rho_\lambda) + I_K(\rho_\lambda) \leq J(0). \tag{6.15}$$

Sine J and I_K are nonnegative functions, we have

$$\lambda TV(\rho_\lambda) \leq J(0). \tag{6.16}$$

If $J(0) = 0$, then $TV(\rho_\lambda) = 0$ for all $\lambda > 0$. If $J(0) > 0$, $TV(\rho_\lambda) \leq \dfrac{J(0)}{\lambda}$. In both cases, $TV(\rho_\lambda) \underset{\lambda \to +\infty}{\longrightarrow} 0$, that is to say $T(\lambda) \underset{\lambda \to +\infty}{\longrightarrow} 0$.

Now consider the continuity of T. Introduce $(\lambda_n)_{n\in\mathbb{N}}$ a sequence of elements of $\mathbb{R}_+$ that converges to λ. In what follows, it is proved that ρ_{λ_n} weakly converges toward ρ_λ. From the definition of ρ_{λ_n}, we have, for all $\rho \in H_0 \cap BV(Q)$ and $n > 0$,

$$J(\rho_{\lambda_n}) + \lambda_n TV(\rho_{\lambda_n}) + I_K(\rho_{\lambda_n}) \leq J(\rho) + \lambda_n TV(\rho) + I_K(\rho). \tag{6.17}$$

The sequence $(J(\rho)+\lambda_n TV(\rho))_{n\in\mathbb{N}}$ converges to $(J(\rho)+\lambda TV(\rho))$, so the sequence $(J(\rho_n))_n$ is bounded, and then $(\|\rho_{\lambda_n}\|_{H(v,Q)})_{n\in\mathbb{N}}$ is bounded as well. This implies that it exists $\rho^* \in H(v,Q)$ such that $(\rho_{\lambda_n})_{n\in\mathbb{N}}$ weakly converges in $H(v,Q)$ and therefore in $L^2(Q)$) to ρ^*. The sequence $(\rho_{\lambda_n})_{n\in\mathbb{N}}$ is bounded in $L^2(Q)$ and then in $L^1(Q)$, since Q is bounded. Moreover, $(TV(\rho_{\lambda_n}))_{n\in\mathbb{N}}$ is also bounded. Consequently $(\rho_{\lambda_n})_{n\in\mathbb{N}}$ is bounded in $BV(Q)$. Accounting for compact embedding results, we get $(\rho_{\lambda_n})_{n\in\mathbb{N}}$ strongly converges to $\rho^* \in BV(Q)$ in $L^1(Q)$, up to a subsequence.
Since $\rho \mapsto \left\| \left(\widetilde{v} \,|\, \widetilde{\nabla}\rho \right) - f \right\|_{L^2(Q)}$ is convex and l.s.c., we have

$$J(\rho^*) \leq \liminf_{n\to+\infty} J(\rho_{\lambda_n}), \tag{6.18}$$

and, the l.s.c. property of the total variation gives

$$TV(\rho^*) \leq \liminf_{n\to+\infty} TV(\rho_{\lambda_n}). \tag{6.19}$$

Accounting for the convexity of K, since it is a closed, it is also weakly closed. Then, the function I_K is l.s.c., and

$$I_K(\rho^*) \leq \liminf_{n\to+\infty} I_K(\rho_{\lambda_n}). \tag{6.20}$$

In the same way, we have, Ker$\gamma_{\widetilde{n}_-}$ is weakly closed in $H(v,Q)$, then we deduce that $\rho^* \in H_0 \cap BV(Q)$. Finally, we have, for all $\rho \in H_0 \cap BV(Q)$,

$$\begin{aligned}
J(\rho^*) + \lambda TV(\rho^*) + I_K(\rho^*) &\leq \liminf_{n\to+\infty} J(\rho_{\lambda_n}) + (\liminf_{n\to+\infty} \lambda_n)(\liminf_{n\to+\infty} TV(\rho_{\lambda_n})) \\
&\quad + \liminf_{n\to+\infty} I_K(\rho_{\lambda_n}) \\
&\leq \liminf_{n\to+\infty} \left(J(\rho_{\lambda_n}) + \lambda_n TV(\rho_{\lambda_n}) + I_K(\rho_{\lambda_n})\right) \\
&\leq J(\rho) + \lambda TV(\rho) + I_K(\rho),
\end{aligned}$$

so $\rho^* = \rho_\lambda$. Moreover, it is clear that ρ_λ is the unique weak limit point of the sequence $(\rho_{\lambda_n})_{n\in\mathbb{N}}$.
Assume $T(\lambda_n) = TV(\rho_{\lambda_n}) \underset{n\to+\infty}{\longrightarrow} t$, first it is proved that T has a closed graph.

$$\begin{aligned}
J(\rho_\lambda) + \lambda t + I_K(\rho_\lambda) &\leq \liminf_{n\to+\infty} J(\rho_{\lambda_n}) + \lambda t + \liminf_{n\to+\infty} I_K(\rho_{\lambda_n}) \\
&= \liminf_{n\to+\infty} J(\rho_{\lambda_n}) + (\liminf_{n\to+\infty} \lambda_n)(\liminf_{n\to+\infty} TV(\rho_{\lambda_n})) \\
&\quad + \liminf_{n\to+\infty} I_K(\rho_{\lambda_n}) \\
&\leq \liminf_{n\to+\infty} \left(J(\rho_{\lambda_n}) + \lambda_n TV(\rho_{\lambda_n}) + I_K(\rho_{\lambda_n})\right) \\
&\leq J(\rho_\lambda) + \lambda TV(\rho_\lambda) + I_K(\rho_\lambda).
\end{aligned}$$

For $\lambda > 0$ fixed, this gives $t \leq TV(\rho_\lambda)$. If $\lambda = 0$, the same inequality holds true. From the decreasing property of T we deduce:

$$TV(\rho_{\lambda_n}) \leq TV(\rho_0), \text{ for all } n > 0. \tag{6.21}$$

Thus, for all $\lambda \geq 0$, we have

$$t \leq TV(\rho_\lambda). \tag{6.22}$$

Gather the previous inequality, with (6.19), we have

$$t = TV(\rho_\lambda), \tag{6.23}$$

which proves that T has a closed graph. Since T is a decreasing function, T is bounded and we conclude that T is continuous.
Our concern is to prove the continuity of Y when λ_n converges to λ, by doing the same way as before. Introduce $S : \underset{\lambda}{\mathbb{R}_+} \to \underset{\mapsto J(\rho_\lambda)}{\mathbb{R}_+}$. If $S(\lambda_n) = J(\rho_{\lambda_n}) \underset{n\to+\infty}{\longrightarrow} s$,

$$\begin{aligned}
s + \lambda TV(\rho_\lambda) + I_K(\rho_\lambda) &= \liminf_{n\to+\infty} J(\rho_{\lambda_n}) + (\liminf_{n\to+\infty} \lambda_n) TV(\rho_\lambda) + I_K(\rho_\lambda) \\
&\leq \liminf_{n\to+\infty} J(\rho_{\lambda_n}) + (\liminf_{n\to+\infty} \lambda_n)(\liminf_{n\to+\infty} TV(\rho_{\lambda_n})) \\
&\quad + \liminf_{n\to+\infty} I_K(\rho_{\lambda_n}) \\
&\leq \liminf_{n\to+\infty} \left(J(\rho_{\lambda_n}) + \lambda_n TV(\rho_{\lambda_n}) + I_K(\rho_{\lambda_n})\right) \\
&\leq J(\rho_\lambda) + \lambda TV(\rho_\lambda) + I_K(\rho_\lambda).
\end{aligned}$$

For all $\lambda \geq 0$,

$$s \leq J(\rho_\lambda), \tag{6.24}$$

and then with (6.18),

$$s = J(\rho_\lambda). \tag{6.25}$$

This shows that the graph of S is closed. Moreover, (6.15) implies that S is bounded. Then, S is continuous.
Finally, let us prove the continuity of Y. We recall that ρ_λ is the unique weak limit point of $(\rho_{\lambda_n})_{n\in\mathbb{N}}$. Moreover, the continuity of S yields that $\|\rho_{\lambda_n}\|_{H(v,Q)} \underset{n\to+\infty}{\longrightarrow} \|\rho_\lambda\|_{H(v,Q)}$. Consequently, because $H(v, Q)$ is a Hilbert space, the subsequence of $(\rho_{\lambda_n})_{n\in\mathbb{N}}$ that weakly converges to ρ_λ strongly converges to ρ_λ as well. This shows that ρ_λ is the unique limit point of $(\rho_{\lambda_n})_{n\in\mathbb{N}}$. We conclude that $\rho_{\lambda_{n\in\mathbb{N}}}$ strongly converges to ρ_λ. The continuity of Y follows.
□

LEMMA 6.4. *Let ρ_0 be the solution to the problem 6.3 for $\lambda = 0$, and assume that $TV(\rho_0) > 0$. For all $\tau \in (0, TV(\rho_0))$, There exists $\lambda \in \mathbb{R}_+$ and $\rho_\lambda \in H_0 \cap BV(Q)$ solution to the problem 6.3 such that $TV(\rho_\lambda) = \tau$.*
Proof.
Let $0 < \tau \in (0, TV(\rho_0))$ be given, Lemma 6.3 claims the existence of $0 < \mu$ sufficiently large such that $T(\mu) - \tau < 0$. Accounting for $0 < T(0) - \tau$, the continuity of function T provides the existence of $0 < \lambda < \mu$ such that $T(\lambda) = TV(\rho_\lambda) = \tau$. Note that a dichotomy process with respect to the parameter λ can be used for finding a non negative solution to the transport equation, in the least squares sense with a given total variation.
□

6.1. A Finite element least squares formulation with penalized total variation and non negativity. In theis section in order to keep the notations as simple as possible we consider only the 1D case. The 2D or 3D situations are directly deduced. Keeping the same notations as in section 4, introduce the problem (6.3) approximated with a Lagrange's finite element method. For $0 \leq \lambda$ be fixed, it reads:

$$u_h = \underset{c_h \in V_h}{\operatorname{argmin}} \quad J(c_h) + \lambda TV(c_h) + I_{K_h}(c_h), \tag{6.26}$$

where $TV(c_h) = \left\| \widetilde{\nabla} c_h \right\|_{L^1(Q)}$.
For numerical convenience (because the transposed operator of the nabla operator is not the divergence operator in finite element context), we replace the total variation by the following discrete operator :

$$TV_d(u) = \sum_{i=1}^{N} \left\| \widetilde{\nabla}_d u(a_i) \right\|_2,$$

where the meshed domain is supposed to be composed of N nodes $a_i = (x_i, y_i)$, $i = 1, \ldots, N$,
with

$$\widetilde{\nabla}_d u(a_i) = \left(\nabla_d^1 u(a_i), \nabla_d^2 u(a_i) \right)^t, \tag{6.27}$$

and

$$\| \nabla_d u(a_i) \|_2 = \sqrt{\left(\nabla_d^1 u(a_i) \right)^2 + \left(\nabla_d^2 u(a_i) \right)^2}. \tag{6.28}$$

The discrete operators are defined by:

$$\begin{aligned} \nabla_d^1 u(a_i) &= \begin{cases} \frac{1}{h_x} \left(u(x_i + h_x, y_i) - u(x_i, y_i) \right) & \text{if } x_i < N_1 \\ 0 & \text{if } x_i = N_1, \end{cases} \\ &= \begin{cases} \frac{1}{h_x} \left(u(a_{i+1}) - u(a_i) \right) & \text{if } i \not\equiv 0 \, [N_1] \\ 0 & \text{if } i \equiv 0 \, [N_1] \end{cases} \end{aligned} \tag{6.29}$$

and

$$\begin{aligned} \nabla_d^2 u(a_i) &= \begin{cases} \frac{1}{h_y} \left(u(x_i, y_i + h_y) - u(x_i, y_i) \right) & \text{if } y_i < N_2 \\ 0 & \text{if } y_i = N_2. \end{cases} \\ &= \begin{cases} \frac{1}{h_y} \left(u(a_{i+N_1}) - u(a_i) \right) & \text{if } i < N_2 \\ 0 & \text{if } i = N_2 \end{cases}, \end{aligned} \tag{6.30}$$

with N_1 is the number of rows and N_2 the number of columns of the mesh. So the discrete problem to solve reads: for $0 < \lambda$ given find:

$$u_h = \underset{c_h \in V}{\operatorname{argmin}} \quad J(c_h) + \lambda TV_d(c_h) + \mathbf{1}_K(c_h). \tag{6.31}$$

Now consider a matricial formulation of the problem (6.31). Write $u_h = \sum_{k=1}^{N} U_k \varphi_k$, with $U \in \mathbb{R}^N$. We have: $\widetilde{\nabla}_d U = (\widetilde{\nabla}_d u(a_1), \cdots \widetilde{\nabla}_d u(a_N))^t$, and keeping the same notations as in section 4, define the functions

$$g_1(U) = \frac{1}{2} U^\top A U - U^\top B; \quad g_2(U) = \lambda TV_d(U) + I_{\mathbb{R}_+^N}(U). \tag{6.32}$$

For $0 \leq \lambda$ be fixed, the descrite problem reads:

$$U = \underset{V \in \mathbb{R}^N}{\operatorname{argmin}} \quad g_1(V) + g_2(V). \tag{6.33}$$

The function g_1 is convex and continuously differentiable, the function g_2 is convex because it is the sum of a continuous convex function (TV_d) and of the indicator function of a closed convex set ($\mathbb{R}_+^N$). Thus we have.
THEOREM 6.5. *Assume that $\lambda \geq 0$. the problem (6.33) has a unique solution.*

Let us mention that a mixed $\mathbb{P}_1/\mathbb{P}_0$ finite element schemes has been proposed for computing the $TV(c_h)$ in [8]. Unfortunately, this method does not extend to other families of finite element.

6.2. Algorithms for computing the solution. This section is dedicated to algorithms for computing a solution to the problem (6.33). Since the function g_1 is a Lipschitz function, the constant of which is denoted by L, the FISTA algorithm introduced in [10] can be used. It reads:

ALGORITHM 6.1.
1. Set $V_1 = U^0 \in \mathbb{R}^N$, and $t_1 = 1$.
2. For U^n and t^n given compute up to convergence :

$$- \ V^n = \underset{X \in \mathbb{R}^N}{\arg\min} \left\{ g_2(X) + \frac{L}{2} \left\| X - \left(U^n - \frac{1}{L} \nabla g_1(U^n) \right) \right\|_2^2 \right\}, \tag{6.34}$$

$$- \ t_{n+1} = \frac{1 + \sqrt{1 + 4t_n^2}}{2}, \tag{6.35}$$

$$- \ U^{n+1} = V^n + \left(\frac{t_n - 1}{t_{n+1}} \right) \left(V^n - V^{n-1} \right). \tag{6.36}$$

Now let us specify how to compute the solution to the equation (6.34). Set

$$\overline{U} = U^n - \frac{1}{L} \nabla g_1(U^n),$$

the problem (6.34) reads

$$U = \underset{X \in \mathbb{R}^N}{\arg\min} \left\{ \frac{L}{2\lambda} \left\| X - \overline{U} \right\|_2^2 + TV_d(X) + I_{\mathbb{R}_+^N}(X) \right\}. \tag{6.37}$$

Define the functions

$$G(U) = \frac{L}{2\lambda} \left\| U - \overline{U} \right\|_2^2 + I_{\mathbb{R}_+^N}(U),$$

and

$$F(\widetilde{\nabla}_d U) = \left\| \widetilde{\nabla}_d U \right\|_2 = TV_d(U).$$

Thanks to the term $\left\|U-\overline{U}\right\|_2^2$, G is 1-uniformly convex, and, since F is convex and continue, we know that, for all $P \in \mathbb{R}^N \times \mathbb{R}^N$

$$F(P)=\sup_{P^* \in \mathbb{R}^N \times \mathbb{R}^N} \langle P^*, P\rangle - F^*(P^*), \tag{6.38}$$

where F^* is the Legendre's transform of F. Moreover, it is well known that $F^* = I_{\mathcal{B}}$, where

$$\mathcal{B}=\left\{P \in \mathbb{R}^K \times \mathbb{R}^K \mid \|P\|_\infty \leq 1\right\}.$$

So we have,

$$F(\nabla_d U)=\sup_{P \in \mathbb{R}^K \times \mathbb{R}^K} \langle P, \nabla_d U\rangle - I_{\mathcal{B}}(P).$$

The following primal-dual formulation of problem (6.37) is deduced:

$$\inf_{U \in \mathbb{R}^K} \sup_{P \in \mathbb{R}^K \times \mathbb{R}^K} \langle P, \nabla_d U\rangle + \frac{L}{2\lambda}\left\|U-\overline{U}\right\|_2^2 + I_{\mathbb{R}_+^N}(U) - I_{\mathcal{B}}(P). \tag{6.39}$$

To solve the saddle point problem (6.39), the following algorithm proposed in [12] is used.

ALGORITHM 6.2.

- *For* $(X^0, Y^0) \in \mathbb{R}^N \times \left(\mathbb{R}^N \times \mathbb{R}^N\right)$ *be fixed and* $\tau_0, \sigma_0 > 0$*, so that* $\tau_0\sigma_0 \left\|\widetilde{\nabla}_d X^0\right\|^2 < 1$ *be fixed, set* $\overline{X}^0 = X^0$.
- *For* $(n \geq 0)$*, update* $X^n, Y^n, \overline{X}^n, \theta_n, \tau_n, \sigma_n$ *as follow :*

$$\begin{cases} Y^{n+1} = \mathrm{prox}_{\sigma F^*}(Y^n + \sigma \widetilde{\nabla}_d \overline{X}^n), \\ X^{n+1} = \mathrm{prox}_{\tau G}(X^n - \tau \widetilde{\nabla}_d^* Y^{n+1}), \\ \theta_n = \dfrac{1}{\sqrt{1+\tau_n}},\ \tau_{n+1} = \theta_n \tau_n,\ \sigma_{n+1} = \dfrac{\sigma_n}{\theta_n}, \\ \overline{X}^{n+1} = X^{n+1} + \theta_n (X^{n+1} - X^n), \end{cases}$$

  where $\widetilde{\nabla}_d^*$ *is the adjoint operator of* $\widetilde{\nabla}_d$.

In [12], the following convergence result is proved.

LEMMA 6.6. *For F a convex, proper, l.s.c. and G a 1-uniformly convex, proper, l.s.c., the sequence (X_n, Y_n) defined in algorithm 6.2 converges to the solution of (6.39).*

Now, let us specify the proximal projections.

LEMMA 6.7.

$$P = \mathrm{prox}_{\sigma F^*}(\widetilde{P}) \Longleftrightarrow P_i = \frac{\widetilde{P}_i}{\max\left(1, \left\|\widetilde{P}_i\right\|_{\mathbb{R}^2}\right)},$$

and

$$U = \mathrm{prox}_{\tau G}(\widetilde{U}) \Longleftrightarrow U_i = \max\left(0, \frac{\lambda \widetilde{U}_i + \tau L \overline{U}_i}{\lambda + \tau L}\right).$$

Proof.
Set $P = \mathrm{prox}_{\sigma F^*}(\widetilde{P})$. Then

$$P = \arg\min_X \left\{ \frac{\left\| X - \widetilde{P} \right\|_2^2}{2\sigma} + F^*(X) \right\}.$$

This yields that $\widetilde{P} \in (I + \sigma \partial F^*)(P)$, that is to say $\dfrac{\widetilde{P} - P}{\sigma} \in \partial F^*(P)$. Since $F^* = I_{\mathcal{B}}$ with $\mathcal{B}$ a closed, non empty convex set, and $\sigma \geq 0$, $\dfrac{\widetilde{P} - P}{\sigma}$ is characterized by

$$\forall V \in \mathcal{B}, \quad \left(\frac{\widetilde{P} - P}{\sigma} \mid V - P \right) \leq 0,$$

so that P is a L^2 projection, $P = \Pi_{\mathcal{B}}(\widetilde{P})$, which proves the first equality. To prove the second one, set $U = \mathrm{prox}_{\tau G}(\widetilde{U})$. Then

$$u = \arg\min_X \left\{ \frac{\left\| X - \widetilde{U} \right\|_2^2}{2\tau} + G(X) \right\}.$$

The first optimality criterion writes :

$$\frac{1}{\tau}(\widetilde{U} - U) + \frac{L}{\lambda}(\overline{U} - U) \in \partial I_{\mathbb{R}_+^N}(U),$$

which is equivalent to

$$\frac{\lambda \widetilde{U} + \tau L \overline{U}}{\lambda \tau} - \frac{\lambda + \tau L}{\lambda \tau} U \in \partial I_{\mathbb{R}_+^N}(U).$$

Since $\mathbb{R}_+^N$ is a closed convex subset, the following characterization holds true: for all $c > 0$,

$$\forall V \in \mathbb{R}_+^N, \quad \left(U + \frac{1}{c} \left(\frac{\lambda \widetilde{U} + \tau L \overline{U}}{\lambda \tau} - \frac{\lambda + \tau L}{\lambda \tau} U \right) - U \mid V - U \right) \leq 0.$$

For $c = \dfrac{\lambda + \tau L}{\lambda \tau}$, we have

$$\forall V \in C_h, \quad \left(\frac{\lambda \widetilde{U} + \tau L \overline{U}}{\lambda + \tau L} - U \mid V - U \right) \leq 0,$$

that is to say

$$U = \Pi_{\mathbb{R}_+^N} \left(\frac{\lambda \widetilde{U} + \tau L \overline{U}}{\lambda + \tau L} \right).$$

□
Remind

$$\nabla g_1(U) = AU - B,$$

thus the algorithm (6.1) reads:
ALGORITHM 6.3.

- *Initialization : For g_1 be a L-lipschitz function, set $V_1 = U^0 \in \mathbb{R}^N$, and $t_1 = 1$.*

- *For $n = 0$ to $n_{max} - 1$: update $X^n, Y^n, \overline{X}^n$ as follows :*

  1− $U^{n+\frac{1}{2}} = U^n - \dfrac{1}{L}(AU^n - F)$

  2− **a**. *For $\widetilde{L}$ a Lipschitz constant of $\widetilde{\nabla}_d$, choose $\tau_0, \sigma_0 > 0$, so that $\tau_0\sigma_0\widetilde{L}^2 < 1$, $X^0 \in \mathbb{R}^N$, $Y^0 \in \mathbb{R}^N \times \mathbb{R}^N$ and set $\overline{X}^0 = X^0$.*

  b. *For $k = 0$ to $k_{max} - 1$*

  $$Y_i^{k+1} = \frac{Y_i^k + \sigma\widetilde{\nabla}_d\overline{X}_i^k}{\max\left(1, \left\|Y_i^k + \sigma\widetilde{\nabla}_d\overline{X}_i^n\right\|_{\mathbb{R}^2}\right)}, \quad i = 1, \dots, N$$

  $$X_i^{k+1} = \max\left\{0\,,\ \frac{\lambda X_i^k - \tau\lambda\widetilde{\nabla}_d^* Y_i^{k+1} + \tau L\, U_i^{n+\frac{1}{2}}}{\lambda + \tau L}\right\}, \quad i = 1, \dots, N$$

  $$\theta_k = \frac{1}{\sqrt{1 + \tau_k}}, \ \tau_{k+1} = \theta_k\tau_k, \ \sigma_{k+1} = \frac{\sigma_k}{\theta_k}$$

  $$\overline{X}^{k+1} = X^{k+1} + \theta_k(X^{k+1} - X^k)$$

  end

  3− $V^n = \overline{X}^{k_{max}}$

  4− $t_{n+1} = \dfrac{1 + \sqrt{1 + 4t_n^2}}{2}$

  5− $U^{n+1} = V^n + \left(\dfrac{t_n - 1}{t_{n+1}}\right)\left(V^n - V^{n-1}\right)$

  end

6.3. Numerical results. Let us end this section with some numerical results. First, for the 1D example with a monotonic initial condition investigated in the previous sections. In figure 6.1 , the solution, computed with a least squares marching technique after 20 time step is represented on the left. In the middle, the positivity constraint has been added and on the right, the positivity constraint and a penalized total variation with $\lambda = 2$.
In the next figure 6.2 the L^2 error is plotted in logarithmic scale for the least squares marching formulation in red, and in blue, for the formulation where a positivity and a total variation constraints has been added. The order of convergence is 2 for both.

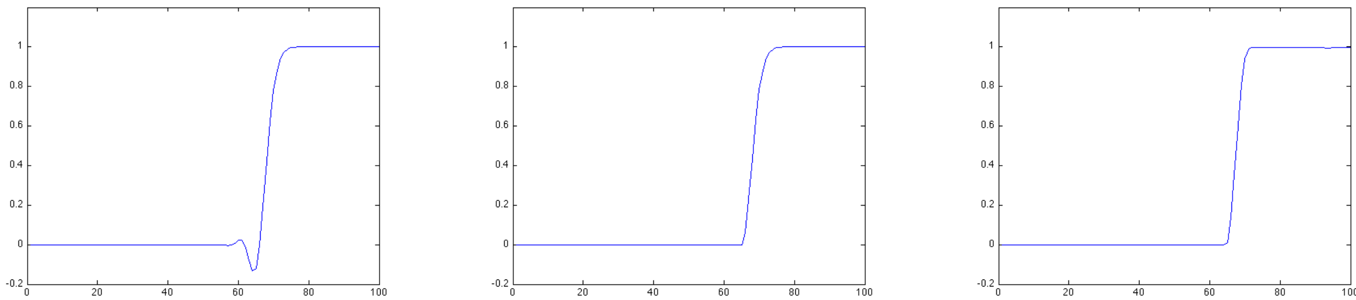


FIG. 6.1. *Left least squares marching solution; middle with positivity constraint; left with positivity and TV constraint.*

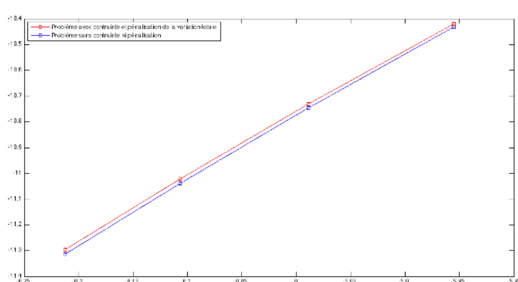


FIG. 6.2. *Convergence curves for LS formulation and for the LS formulation with positivity and TV constraints.*

A 2D example is now considered. A bump subject to a rotating velocity field. In figure 6.3 the velocity field and the initial position of the bump are given.

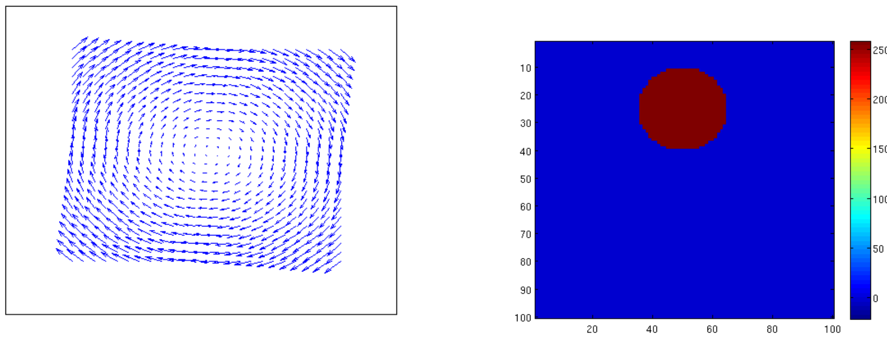


FIG. 6.3. *left velocity field, right initial position of the bump*

In figure 6.4 results for the LS formulation are presented. On first row, on left, a projection on the x, y plan of the bump is plotted, on right, a slice on an horizontal plan going through the center of the bump is represented. On the second row, on left, a slice on a vertical plan going through the center of the bump is represented, on right, the bump is represented after a rotation of $\frac{\pi}{2}$. In figure 6.5 results for the LS formulation with a positivity constraint are

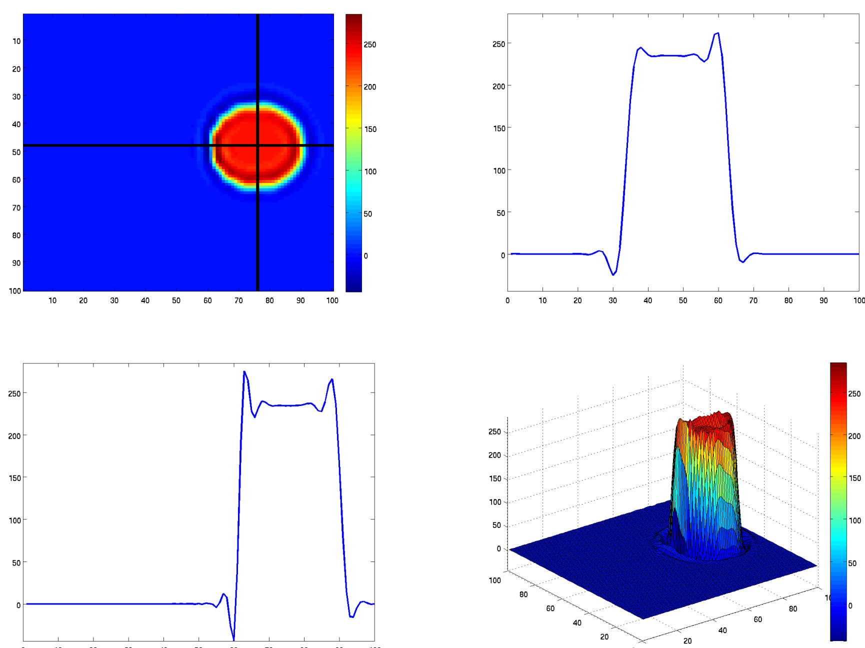


FIG. 6.4. *LS formulation*

presented. On first row, on left, a projection on the x, y plan of the bump is plotted, on right a slice on an vertical plan going through the center of the bump is represented. On the second row, on left, a slice on a horizontal plan going through the center of the bump is represented, on right, the bump is represented after a rotation of $\frac{\pi}{2}$.

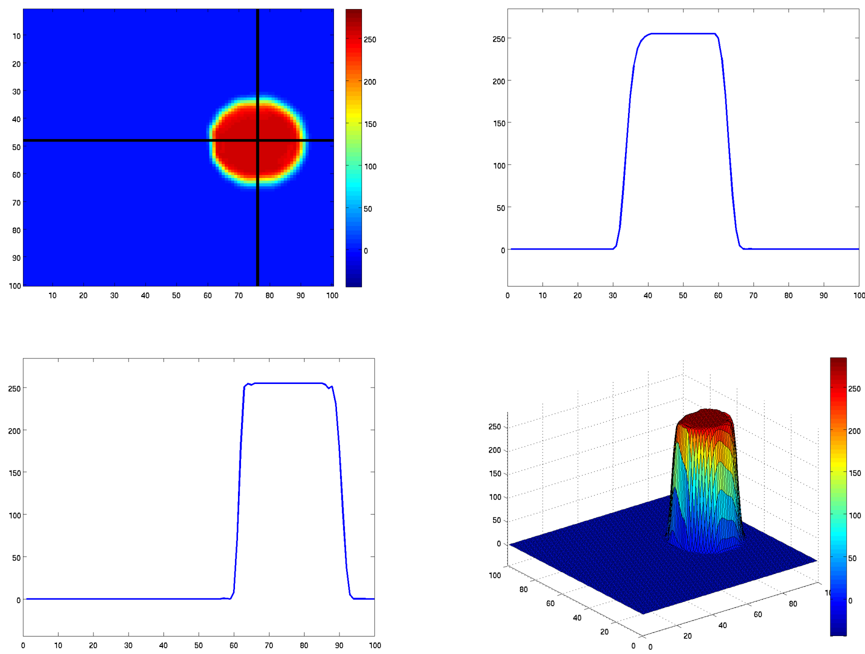


FIG. 6.5. *LS formulation with positivity constraint*

In figure 6.6 results for the LS formulation with positivity and TV constraints for $\lambda = 2$ are presented. On first row, on left, a projection on the x, y plan of the bump is plotted, on right a slice on an vertical plan going through the center of the bump is represented. On the second row, on left, a slice on a horizontal plan going through the center of the bump is represented, on right, the bump is represented after a rotation of $\frac{\pi}{2}$.
Let us come back to the non monotonic case. In figure 6.3, on the left, the least squares marching solution is presented, on the middle, the least squares marching with a positivity constraint solution, and on the lright, the least squares marching with positivity and TV constraints solution. It can be checked that spurious oscillations desappear with positivity and TV constraints.
A 2D example is now considered with a non monotonic bump subject to a rotating velocity field as in previous case. In figure 6.7 the velocity field and the initial position of the bump are given.
In figure 6.8 results for the LS formulation are presented. On first row, on left, a projection on the x, y plan of the bump is plotted, on right, a slice on an horizontal plan going through the center of the bump is represented. On the second row, on left, a slice on a vertical plan going through the center of the bump is represented, on right, the bump is represented after a rotation of $\frac{\pi}{2}$. The numerical scheme damps a part of the oscillations in the direction of the velocity field, but not in the orthogonal direction.
In figure 6.9 results for the LS formulation with a positivity constraint are presented. On first row, on left, a projection on the x, y plan of the bump is plotted, on right, a slice on an horizontal plan going through the center of the bump is represented. On the second row, on

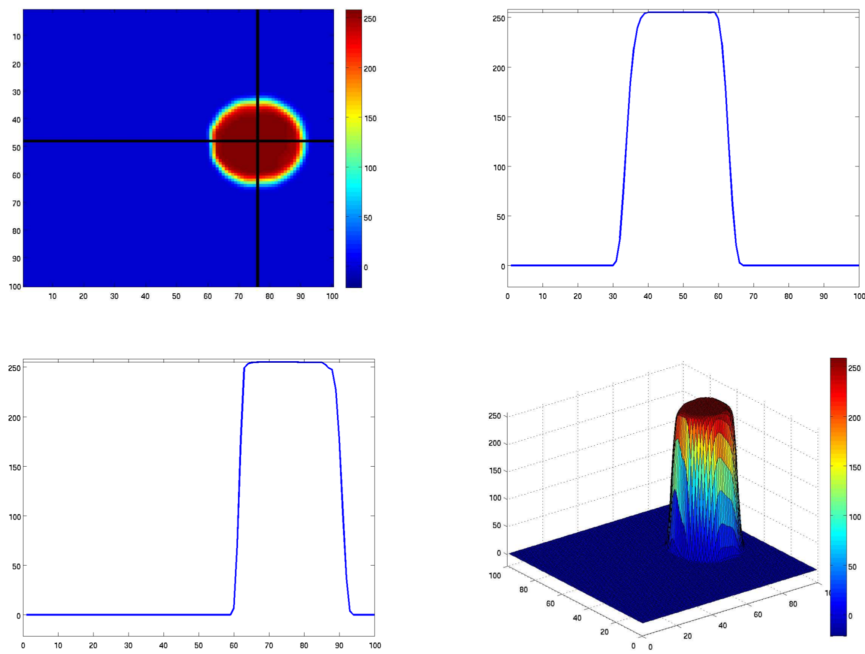


FIG. 6.6.

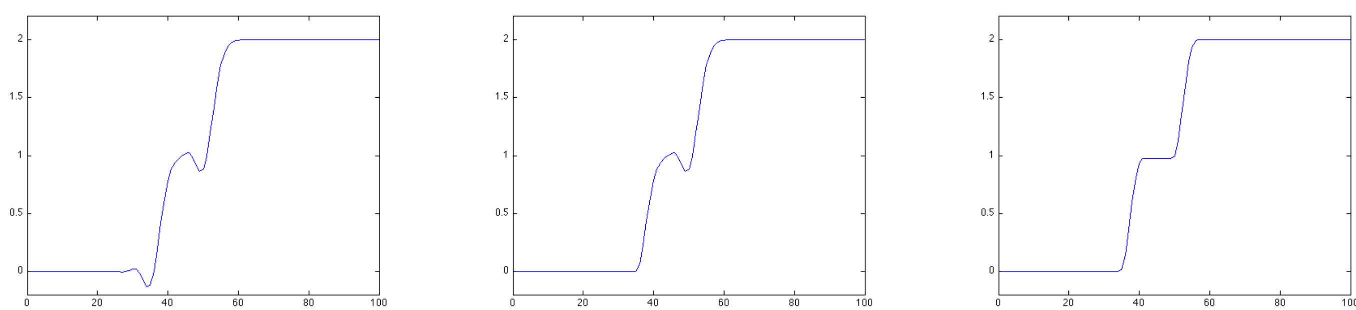


FIG. 6.7.

left, a slice on a vertical plan going through the center of the bump is represented, on right, the bump is represented after a rotation of $\frac{\pi}{2}$. The numerical scheme with a positivity constraint damps the oscillations in the direction of the velocity field, but not in the orthogonal direction. In figure 6.10 results for the LS formulation with positivity and TV constraints with $\lambda = 2$ are presented. On first row, on left, a projection on the x, y plan of the bump is plotted, on right, a slice on an horizontal plan going through the center of the bump is represented. On the second row, on left, a slice on a vertical plan going through the center of the bump is represented, on right, the bump is represented after a rotation of $\frac{\pi}{2}$. The numerical scheme with a positivity and TV constraints handle the oscillations in both directions. In the last figure 6.11, the influence of lambda is investigated with a positivity constraint. In the following table the values of the total variation of solutions are reported.

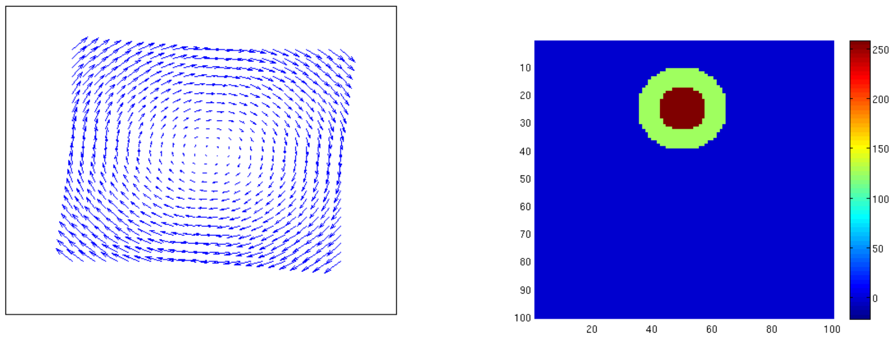


FIG. 6.8.

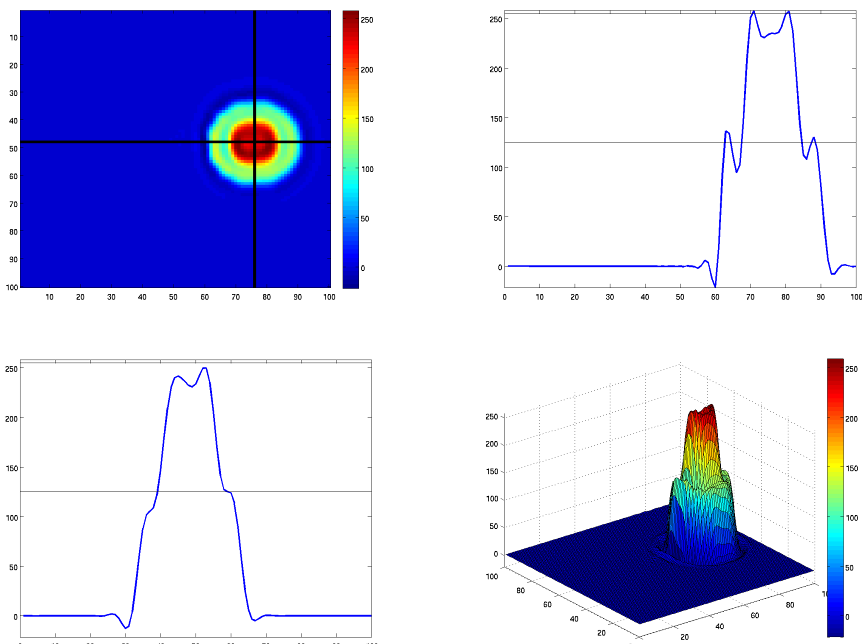


FIG. 6.9.

Solution	TV	Postivity constraint	TV constraint λ
initial condition	2.0710^4	no	no $\lambda = 0$
LS marching	2.2710^4	no	no $\lambda = 0$
LS marching	1.9310^4	yes	no $\lambda = 0$
LS marching	1.7410^4	yes	yes $\lambda = 2$

We can conclude that the proposed numerical scheme which control the extreme values of the solution (through indicator functions) with a penalization of its total variation is able to cancel the spurious oscillations whatever the initial condition is. Moreover, this numerical scheme does not affect the convergence order of the lagrange finite element method. The proposed method is an acceptable answer to the deficiency of the finite element method for

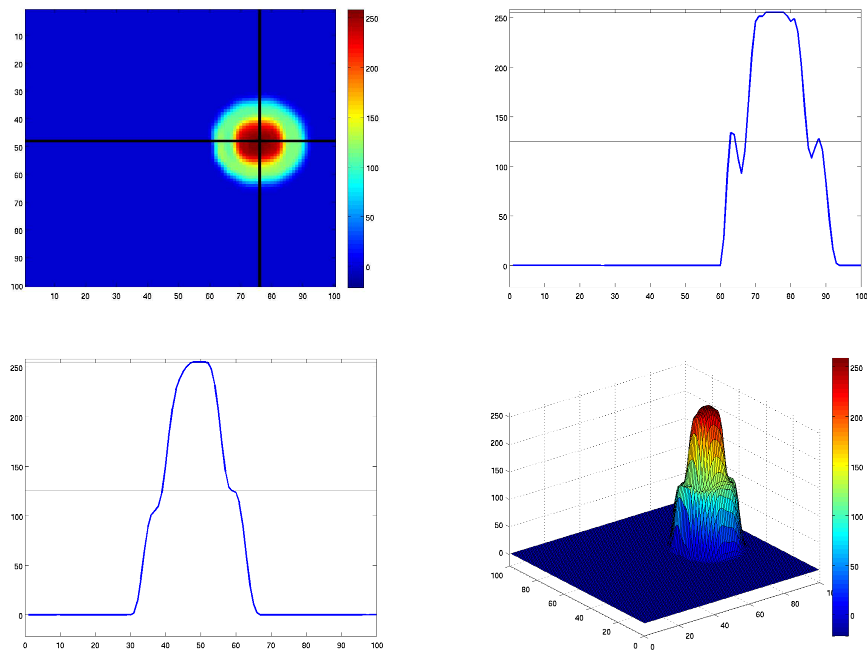


FIG. 6.10.

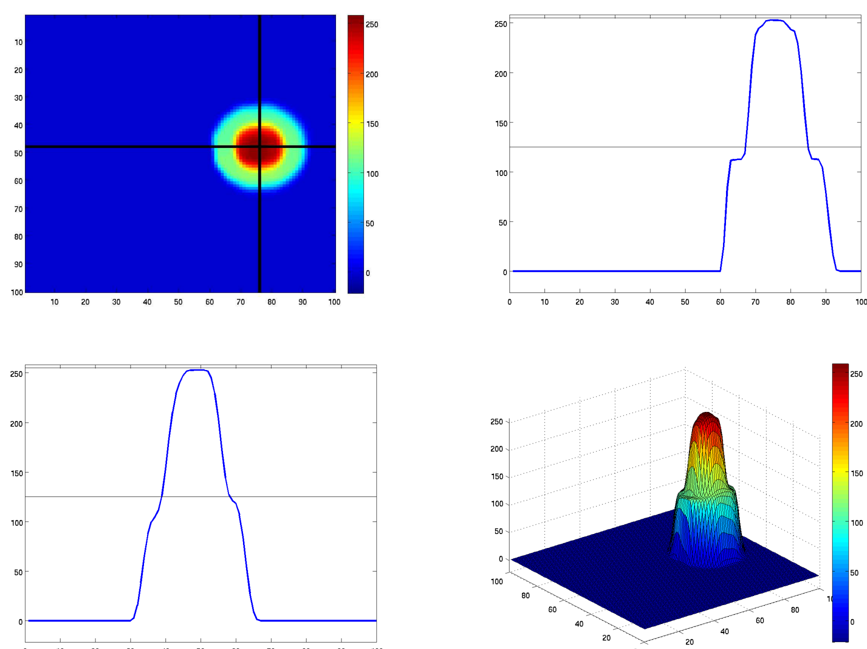
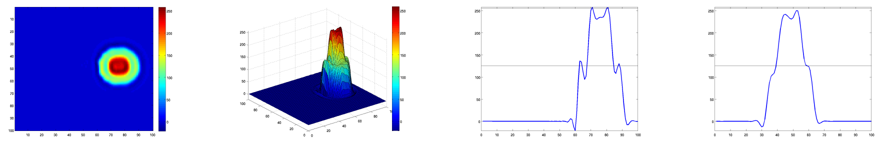
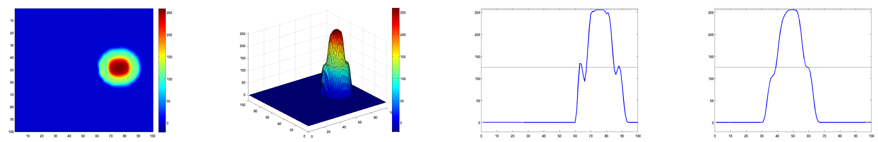


FIG. 6.11.

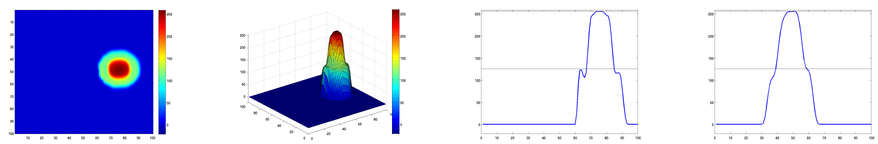
$\lambda = 0$



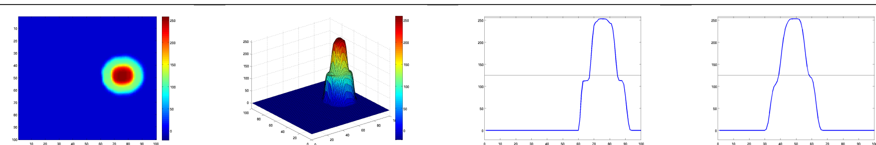
$\lambda = 0$ with MP



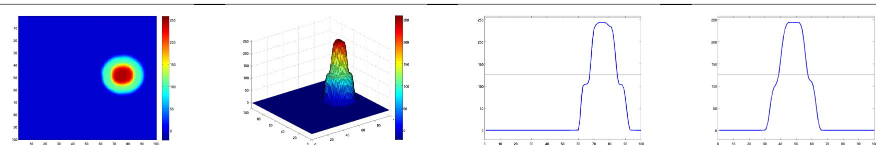
$\lambda = 1$ with MP



$\lambda = 2$ with MP



$\lambda = 5$ with MP



$\lambda = 10$ with MP

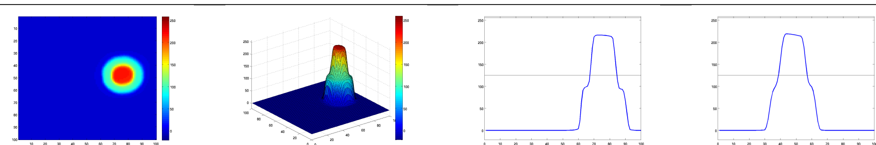


FIG. 6.12. *Influence of the total variation*

handling the transport equation.

Acknowledgments. This work was supported by the LABEX MILYON (ANR-10-LABX-0070) of Universit de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR); and by ANR 11-TecSan002-03.

REFERENCES

[1] AMBROSIO, L., FUSCO, N., PALLARA, D.: Functions of bounded variation and free discontinuity problems.

Oxford mathematical monographs, Oxford University Press (2000).
[2] ATTOUCH, H., BUTTAZZO, MICHAILLE, G.: Variational analysis in Sobolev and BV spaces : applications to PDEs and optimization. MPS-SIAM series on optimization, 2006
[3] AUBERT, G., DERICHE,.R, KORNPROBST, P.: Computing optical flow problem via variational techniques. SIAM J. Appl. Math., 80, 156–182, (1999).
[4] Aubert, G., Kornprobst, P.: Mathematical Problems in Image Processing, Partial Differential Equations and the Calculus of Variations. Applied Mathematical Sciences 147, Springer Verlag (2006).
[5] AZÉRAD, P.: Analyse des équations de Navier-Stokes en bassin peu profond et de l'équation de transport. Thèse de doctorat, Université de Neuchâtel (1996).
[6] AZÉRAD, P.; POUSIN, J.: Inégalité de Poincaré courbe pour le traitement variationnel de l'équation de transport. C. R. Acad. Sci., Paris, Ser. I **322**, 721-727, (1996).
[7] AZÉRAD, P.; PERROCHET, P.; POUSIN, J.: Space-time integrated least-squares: A simple, stable and precise finite element scheme to solve advection equations as if they were *elliptic*. Chipot, M. (ed.) et al., Progress in partial differential equations: the Metz surveys 4. Proceedings of the conference given at the University of Metz, France during the 1994-95 'Metz Days'. Harlow: Longman. Pitman Res. Notes Math. Ser. 345, 161-174 (1996).
[8] BARTELS, S.: Total Variation Minimization with Finite Elements: Convergence and Iterative Solution, SIAM J. Numerical Analysis, **3**, 1162-1180, (2012).
[9] BESSON, O.; DE MONTMOLLIN, G.: A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems, Int. J. Meth. Fluids, **44**, 525-543, (2004).
[10] BECK, A., TEBOULLE, M.: Space-time integrated least squares: a time marching approach. SIAM J. Img. Sci., **1**, 183–202, (2009).
[11] P.B. Bochev and M.D. Gunzburger, Least-Squares Finite Element Methods, volume 166, Applied Mathematical Sciences, Springer, 2009.
[12] CHAMBOLLE, A., POCK, T.: A First-Order Primal-Dual Algorithm for Convex Problemswith Applications to Imaging, Journal of Mathematical Imaging and Vision, **1**, 120-145, (2011).
[13] CHEN, G.-Q.; FRID, H.: Divergence-Measure fields and hyperbolic conservation laws. Arch. Rational Mech. Anal. **147**, 89-118, (1999).
[14] DE MONTMOLLIN, G.: Méthode STILS pour l'équation de transport: comparaisons et analyses. Etude d'un modèle de fermeture pour la loi de Darcy. Thèse de doctorat, Université de Neuchâtel (2001).
[15] DE STERCK H., THOMAS A., MANTEUFFEL A., MACCORMIK STEPHEN F., OLSON L. : Least-squares finite element methods and algebraic multigrid solvers for linear hyperbolic PDEs. SIAM J. Sci Comput. Vol 26, **N 1**, pp. 31-54 (2004).
[16] FRANCA, L.P.; STENBERG, R.: Error analysis of Galerkin least squares methods for the elasticity equations. SIAM-J.-Numer.-Anal., 28 (1991) pp 1680–1697.
[17] GLOWINSKI, R.: Numerical Methods for Nonlinear Variational problems. Springer-Verlag, 1984.
[18] FRANCHI, B,; SERAPIONI, R,; SERRA CASSANO, F,: Meyer-Serrin type theorems and relaxation of variational integrals depending on vectors fields. Houston J. Math. **22**, 859-890, (1996).
[19] KUZMIN, D.:Linearity-preserving flux correction and convergence acceleration for constrained Galerkin schemes. JCAM **236**, (2012) pp. 2317-2337.
[20] LIONS, P.L.: Mathematical topics in fluid mechanics. Vol. 1; Oxford Science Publication, Calderon Press (1996).
[21] BESSON, O,; POUSIN, J,; Solutions for Linear Conservation Laws with Velocity Fields in L^∞. Archive for Rational Mechanics and Analysis, **186**, 159-175, (2007).
[22] ROKAFELLAR AND WETS: Variational analysis (.Springer 2008) Springer (2008).
[23] YOUNES L., ARRATE F., MILLER M.: Evolution Equations in Computational Anatomy. Neuroimage (2008), pp. S40–S50.